

A 5

PCT

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/674,593

DATE: 07/05/2001
TIME: 15:41:02

Input Set : A:\538889_1.txt
Output Set: N:\CRF3\07032001\I674593.raw

3 <110> APPLICANT: Van den Eynde, Benoit
 4 Boon-Falleur, Thierry
 6 <120> TITLE OF INVENTION: TUMOR ASSOCIATED ANTIGEN ENCODED BY THE REVERSE STRAND
 7 OF A NEW UBIQUITOUSLY EXPRESSED GENE
 9 <130> FILE REFERENCE: L0461/7099
 11 <140> CURRENT APPLICATION NUMBER: 09/674,593
 C--> 12 <141> CURRENT FILING DATE: 2000-06-07
 14 <150> PRIOR APPLICATION NUMBER: PCT/US99/10424
 15 <151> PRIOR FILING DATE: 1999-05-13
 17 <150> PRIOR APPLICATION NUMBER: US 60/085,318
 18 <151> PRIOR FILING DATE: 1998-05-13
 20 <160> NUMBER OF SEQ ID NOS: 11
 22 <170> SOFTWARE: FastSEQ for Window Version 3.0
 25 <210> SEQ ID NO: 1
 26 <211> LENGTH: 1382
 27 <212> TYPE: DNA
 28 <213> ORGANISM: Homo sapiens
 30 <220> FEATURE:
 31 <221> NAME/KEY: CDS
 32 <222> LOCATION: 738..989
 34 <400> SEQUENCE: 1
 35 cattatgcta acagcataaaa catgcagggg gtgggagcag ggtcacaaaa gtgagtgtt 60
 36 tcaattctac ttggaatgaa aggttgaat aattaaaca gtacggaaa tgcagagcaa 120
 37 ttttcctc tggtgacaat atagtgtcca acacttggaa gtgattttta agaatgttta 180
 38 tttaaattaa aaggatggat ttccaaaggaa aaaaaataag gaaaaggaaa gaaaaaactg 240
 39 aacagaaaac gcaaaggat cagtttggc actaacctt gcaaggatac ctttttattt 300
 40 tcttaagat tcctttgtt tatacacaga tttaagttt actcctactg ctgaccsaag 360
 41 tgaatttcct tctccagtca cagtgtaac ctctaccccc caactgcaac gagagtttg 420
 42 aggggcatca atcacaccga gaagtacacag cccctcaacc actgaggtgt gggggggtag 480
 43 gatatctgcat ttcttcatat caaccccccaca ctataggcgca cctaaatggg tggcggtgg 540
 44 gggagaccga ctcacttggag tttcttgaag gtttcttggc ctccagccac gtaattgccc 600
 45 ccgcctctgga tctggcttag cttccggatt cggtgccag tccgcgggtt gtagatgttc 660
 46 ctgacggccc caaagggtgc ctgaacgcgcg ccggtcacct ctttcaggaa gacttcgaag 720
 47 ctggacacct tcttc tcttc atg gat gac gac gcg ccc cgc gta gaa ggg 770
 48 Met Asp Asp Asp Ala Ala Pro Arg Val Glu Gly
 49 1 5 10
 50 gtc ccc gtt gcg gta cac aag cac gct ctt cac gac ggg ctg aga cag 818
 51 Val Pro Val Ala Val His Lys His Ala Leu His Asp Gly. Leu Arg Gln
 52 15 20 25
 53 gtg gct gga cct ggc gct gct gcc gct cat ctt ccc cgc tgg ccg ccg 866
 54 Val Ala Gly Pro Gly Ala Ala Ala His Leu Pro Arg Trp Pro Pro
 55 30 35 40
 56 cct cag ctc gct gct tcg cgt cgg gag gca cct ccg ctg tcc cag cgg 914
 57 Pro Gln Leu Ala Ala Ser Arg Arg Glu Ala Pro Pro Leu Ser Gln Arg
 58 45 50 55
 59 cct cac cgc acc cag ggc gcg gga tcg cct cct gaa acg aac gag aaa 962
 60 Pro His Arg Thr Gln Gly Ala Gly Ser Pro Pro Glu Thr Asn Glu Lys

ENTERED

See page 5

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/674,593

DATE: 07/05/2001
TIME: 15:41:02

Input Set : A:\538889_1.txt
Output Set: N:\CRF3\07032001\I674593.raw

61	60	65	70	75			
62	ctg acg aat cca cag gtg aaa gag aag	taacggccgt	gcgccttaggc	gtccacc	1016		
63	Leu Thr Asn Pro Gln Val Lys Glu Lys						
64	80						
65	cagaggagac actaggagct	tgcaggactc	ggagtagacg	ctcaagttt	tcaccgtggc	1076	
66	gtgcacagcc aatcaggacc	cgcagtgcgc	gcaccacacc	aggttcacct	gctacgggca	1136	
67	gaatcaagggt	ggacagcttc	tgagcaggag	ccggaaacgc	gcggggcctt	caaacaggca	1196
68	cgcctagta	gggcaggaga	gaggaggacg	cacacacaca	cacacacaca	aatatggtga	1256
69	aacccaattt	cttacatcat	atctgtgcta	ccctttccaa	acagcctaatt	tttttttttc	1316
70	tctcttcttg	caccttacc	cctcaatctc	ctgcttcctc	ccaaattaaa	gcaattaagt	1376
71	tcctgg					1382	
73	<210> SEQ ID NO: 2						
74	<211> LENGTH: 84						
75	<212> TYPE: PRT						
76	<213> ORGANISM: Homo sapiens						
78	<400> SEQUENCE: 2						
79	Met Asp Asp Asp Ala Ala Pro Arg Val Glu Gly Val Pro Val Ala Val						
80	1 5 10 15						
81	His Lys His Ala Leu His Asp Gly Leu Arg Gln Val Ala Gly Pro Gly						
82	20 25 30						
83	Ala Ala Ala Ala His Leu Pro Arg Trp Pro Pro Pro Gln Leu Ala Ala						
84	35 40 45						
85	Ser Arg Arg Glu Ala Pro Pro Leu Ser Gln Arg Pro His Arg Thr Gln						
86	50 55 60						
87	Gly Ala Gly Ser Pro Pro Glu Thr Asn Glu Lys Leu Thr Asn Pro Gln						
88	65 70 75 80						
89	Val Lys Glu Lys						
91	<210> SEQ ID NO: 3						
92	<211> LENGTH: 9						
93	<212> TYPE: PRT						
94	<213> ORGANISM: Homo sapiens						
96	<400> SEQUENCE: 3						
97	Leu Pro Arg Trp Pro Pro Pro Gln Leu						
98	1 5						
100	<210> SEQ ID NO: 4						
101	<211> LENGTH: 2167						
102	<212> TYPE: DNA						
103	<213> ORGANISM: Homo sapiens						
105	<220> FEATURE:						
106	<221> NAME/KEY: CDS						
107	<222> LOCATION: 303..1730						
109	<400> SEQUENCE: 4						
110	gtccacccatg attctgccccg tagcaggta acctgggtgtg	gtgcgcac	tgccggtcct		60		
111	gattggctgt gcacgccacg	gtgaaaaact	tgagcgctca	ctccgagtcc	tgcaagctcc	120	
112	tagtgtctcc tctgggtgga	cgcctaggcg	cacggccgtt	acttctcttt	cacctgtgga	180	
113	ttcgtcagtt tctcggtcg	ttcaggaggc	gatcccgcgc	cctgggtgcg	gtgaggccgc	240	
114	tgggacagcg	gagggtgcctc	ccgacgcgaa	gcagcgagct	gaggcggcgg	ccagcggggaa	300
115	ag atg agc ggc agc gcc agg tcc agc cac ctg tct	cag ccc gtc					347
116	Met Ser Gly Ser Ser Ala Arg Ser Ser His Leu Ser Gln Pro Val						

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/674,593

DATE: 07/05/2001
TIME: 15:41:02

Input Set : A:\538889_1.txt
Output Set: N:\CRF3\07032001\I674593.raw

117	1	5	10	15	
118	gtg aag agc gtg ctt gtg tac cgc aac ggg gac ccc ttc tac gcg ggg				395
119	Val Lys Ser Val Leu Val Tyr Arg Asn Gly Asp Pro Phe Tyr Ala Gly				
120	20	25	30		
121	cgc cgc gtc gtc atc cat gag aag aag gtg tcc agc ttc gaa gtc ttc				443
122	Arg Arg Val Val Ile His Glu Lys Lys Val Ser Ser Phe Glu Val Phe				
123	35	40	45		
124	ctg aag gag gtg acc ggc ggc gtt cag gca ccc ttt ggg gcc gtc agg				491
125	Leu Lys Glu Val Thr Gly Gly Val Gln Ala Pro Phe Gly Ala Val Arg				
126	50	55	60		
127	aac atc tac acc ccg cg act ggc cac cga atc cgg aag cta gac cag				539
128	Asn Ile Tyr Thr Pro Arg Thr Gly His Arg Ile Arg Lys Leu Asp Gln				
129	65	70	75		
130	atc cag agc ggg ggc aat tac gtg gct gga ggc cag gaa gcc ttc aag				587
131	Ile Gln Ser Gly Gly Asn Tyr Val Ala Gly Gly Gln Glu Ala Phe Lys				
132	80	85	90	95	
133	aaa ctc aat tac ttg gac ata gga gaa atc aag aaa aga cca atg gaa				635
134	Lys Leu Asn Tyr Leu Asp Ile Gly Glu Ile Lys Lys Arg Pro Met Glu				
135	100	105	110		
136	gtt gtt aat aca gag gta aaa cca gta atc cat agc agg atc aac gtg				683
137	Val Val Asn Thr Glu Val Lys Pro Val Ile His Ser Arg Ile Asn Val				
138	115	120	125		
139	tca gct cgc ttt aga aaa ccg ctt cag gag ccg tgc act atc ttc ttg				731
140	Ser Ala Arg Phe Arg Lys Pro Leu Gln Glu Pro Cys Thr Ile Phe Leu				
141	130	135	140		
142	att gca aat gga gac ctc ata aac cca gct tct cgc ctc ctt atc ccc				779
143	Ile Ala Asn Gly Asp Leu Ile Asn Pro Ala Ser Arg Leu Leu Ile Pro				
144	145	150	155		
145	aga aaa acc ttg aat cag tgg gat cat gta cta caa atg gtc aca gaa				827
146	Arg Lys Thr Leu Asn Gln Trp Asp His Val Leu Gln Met Val Thr Glu				
147	160	165	170	175	
148	aaa atc act ctg agg agc ggg gct gtt cac agg ctt tat act tta gaa				875
149	Lys Ile Thr Leu Arg Ser Gly Ala Val His Arg Leu Tyr Thr Leu Glu				
150	180	185	190		
151	gga aaa ctt gtt gag agt gga gca gag ttg gag aat ggg cag ttt tat				923
152	Gly Lys Leu Val Glu Ser Gly Ala Glu Leu Glu Asn Gly Gln Phe Tyr				
153	195	200	205		
154	gtg gct gtt ggc aga gat aag ttt aag aaa ctg cct tac ggt gag tta				971
155	Val Ala Val Gly Arg Asp Lys Phe Lys Lys Leu Pro Tyr Gly Glu Leu				
156	210	215	220		
157	ctt ttt gac aag tca acg atg aga agg cct ttt ggt cag aaa gct tct				1019
158	Leu Phe Asp Lys Ser Thr Met Arg Arg Pro Phe Gly Gln Lys Ala Ser				
159	225	230	235		
160	tca cta cct cct att gta gga tcc aga aag tct aaa ggg agt gga aat				1067
161	Ser Leu Pro Pro Ile Val Gly Ser Arg Lys Ser Lys Gly Ser Gly Asn				
162	240	245	250	255	
163	gat cgc cac tct aag tca aca gtt gga tcc agt gac aac tca tct cct				1115
164	Asp Arg His Ser Lys Ser Thr Val Gly Ser Ser Asp Asn Ser Ser Pro				
165	260	265	270		

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/674,593

DATE: 07/05/2001
TIME: 15:41:02

Input Set : A:\538889_1.txt
Output Set: N:\CRF3\07032001\I674593.raw

166	cag ccc ctg aag agg aaa ggg aaa aaa gaa gac gtg aat tca gaa aaa	1163
167	Gln Pro Leu Lys Arg Lys Gly Lys Lys Glu Asp Val Asn Ser Glu Lys	
168	275 280 285	
169	ctg acg aaa ttg aaa caa aat gta aaa tta aag aat tca caa gaa acc	1211
170	Leu Thr Lys Leu Lys Gln Asn Val Lys Leu Lys Asn Ser Gln Glu Thr	
171	290 295 300	
172	att cca aat agt gat gaa ggc att ttc aaa gct gga gca gag agg tct	1259
173	Ile Pro Asn Ser Asp Glu Gly Ile Phe Lys Ala Gly Ala Glu Arg Ser	
174	305 310 315	
175	gaa aca cgg ggg gca gca gaa gtc caa gaa gat gaa gat act cag gtt	1307
176	Glu Thr Arg Gly Ala Ala Glu Val Gln Glu Asp Glu Asp Thr Gln Val	
177	320 325 330 335	
178	gag gtt cca gtc gat cag agg cca gca gaa ata gta gac gag gaa gaa	1355
179	Glu Val Pro Val Asp Gln Arg Pro Ala Glu Ile Val Asp Glu Glu Glu	
180	340 345 350	
181	gat gga gag aag gca aac aag gat gca gaa cag aaa gaa gac ttt tca	1403
182	Asp Gly Glu Lys Ala Asn Lys Asp Ala Glu Gln Lys Glu Asp Phe Ser	
183	355 360 365	
184	gga atg aat ggt gac ctt gaa gag gaa gga ggt agg gag gct aca gat	1451
185	Gly Met Asn Gly Asp Leu Glu Glu Gly Gly Arg Glu Ala Thr Asp	
186	370 375 380	
187	gcc cct gag caa gtc gag gag att ctg gat cac agt gag cag cag gca	1499
188	Ala Pro Glu Gln Val Glu Glu Ile Leu Asp His Ser Glu Gln Gln Ala	
189	385 390 395	
190	cgc cct gct cgt gta aat gga ggc acc gat gag gag aat ggt gag gag	1547
191	Arg Pro Ala Arg Val Asn Gly Gly Thr Asp Glu Glu Asn Gly Glu	
192	400 405 410 415	
193	ctg cag cag gtt aat aat gag ctt caa ctg gtc cta gac aag gaa aga	1595
194	Leu Gln Gln Val Asn Asn Glu Leu Gln Leu Val Leu Asp Lys Glu Arg	
195	420 425 430	
196	aag tct caa gga gct ggc agt gga caa gat gag gct gat gta gac cct	1643
197	Lys Ser Gln Gly Ala Gly Ser Gly Gln Asp Glu Ala Asp Val Asp Pro	
198	435 440 445	
199	caa aga cca cca agg cca gaa gta aaa att acc agt cca gaa gaa aat	1691
200	Gln Arg Pro Pro Arg Pro Glu Val Lys Ile Thr Ser Pro Glu Glu Asn	
201	450 455 460	
202	gaa aac aac caa caa aac aag gac tat gct gcc gtg gct tagaagattt tt	1742
203	Glu Asn Asn Gln Gln Asn Lys Asp Tyr Ala Ala Val Ala	
204	465 470 475	
205	aaaaagagag tatatggatc gcaagaaaaa tgaagggtta tcataacttga aagataagca	1802
206	catagttatt gctaatata atgtgacact atgtcgaaat actacacctacg aattataaca	1862
207	ttagaaggctt agtggaaaga ccagataact ttaaatggct actaaaggat aattacttac	1922
208	ttttatttgc atgttttaa aagtcatata gaaatattaa ataagacgga cagaggagaa	1982
209	tttgcactgg aagacaattt ccacttgaa agatgaaaaa ataggatcac tcttattgtt	2042
210	cgctttatata taagttttaga aggcaattt ttctaaataa ttttctcta ggaaggcgta	2102
211	gaatttaaa gaactggtaa taggaaagca tgtactattt tcttaaagca ataaacttctt	2162
212	gaatg	2167
214	<210> SEQ ID NO: 5	
215	<211> LENGTH: 476	

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/674,593

DATE: 07/05/2001
TIME: 15:41:02

Input Set : A:\538889_1.txt
Output Set: N:\CRF3\07032001\I674593.raw

216 <212> TYPE: PRT
 217 <213> ORGANISM: Homo sapiens
 219 <400> SEQUENCE: 5
 220 Met Ser Gly Ser Ser Ala Arg Ser Ser His Leu Ser Gln Pro Val Val
 221 1 5 10 15
 222 Lys Ser Val Leu Val Tyr Arg Asn Gly Asp Pro Phe Tyr Ala Gly Arg
 223 20 25 30
 224 Arg Val Val Ile His Glu Lys Val Ser Ser Phe Glu Val Phe Leu
 225 35 40 45
 226 Lys Glu Val Thr Gly Gly Val Gln Ala Pro Phe Gly Ala Val Arg Asn
 227 50 55 60
 228 Ile Tyr Thr Pro Arg Thr Gly His Arg Ile Arg Lys Leu Asp Gln Ile
 229 65 70 75 80
 230 Gln Ser Gly Gly Asn Tyr Val Ala Gly Gly Gln Glu Ala Phe Lys Lys
 231 85 90 95
 232 Leu Asn Tyr Leu Asp Ile Gly Glu Ile Lys Lys Arg Pro Met Glu Val
 233 100 105 110
 234 Val Asn Thr Glu Val Lys Pro Val Ile His Ser Arg Ile Asn Val Ser
 235 115 120 125
 236 Ala Arg Phe Arg Lys Pro Leu Gln Glu Pro Cys Thr Ile Phe Leu Ile
 237 130 135 140
 238 Ala Asn Gly Asp Leu Ile Asn Pro Ala Ser Arg Leu Leu Ile Pro Arg
 239 145 150 155 160
 240 Lys Thr Leu Asn Gln Trp Asp His Val Leu Gln Met Val Thr Glu Lys
 241 165 170 175
 242 Ile Thr Leu Arg Ser Gly Ala Val His Arg Leu Tyr Thr Leu Glu Gly
 243 180 185 190
 244 Lys Leu Val Glu Ser Gly Ala Glu Leu Glu Asn Gly Gln Phe Tyr Val
 245 195 200 205
 246 Ala Val Gly Arg Asp Lys Phe Lys Lys Leu Pro Tyr Gly Glu Leu Leu
 247 210 215 220
 248 Phe Asp Lys Ser Thr Met Arg Arg Pro Phe Gly Gln Lys Ala Ser Ser
 249 225 230 235 240
 250 Leu Pro Pro Ile Val Gly Ser Arg Lys Ser Lys Gly Ser Gly Asn Asp
 251 245 250 255
 252 Arg His Ser Lys Ser Thr Val Gly Ser Ser Asp Asn Ser Ser Pro Gln
 253 260 265 270
 254 Pro Leu Lys Arg Lys Gly Lys Lys Glu Asp Val Asn Ser Glu Lys Leu
 255 275 280 285
 256 Thr Lys Leu Lys Gln Asn Val Lys Leu Lys Asn Ser Gln Glu Thr Ile
 257 290 295 300
 258 Pro Asn Ser Asp Glu Gly Ile Phe Lys Ala Gly Ala Glu Arg Ser Glu
 259 305 310 315 320
 260 Thr Arg Gly Ala Ala Glu Val Gln Glu Asp Glu Asp Thr Gln Val Glu
 261 325 330 335
 262 Val Pro Val Asp Gln Arg Pro Ala Glu Ile Val Asp Glu Glu Glu Asp
 263 340 345 350
 264 Gly Glu Lys Ala Asn Lys Asp Ala Glu Gln Lys Glu Asp Phe Ser Gly
 265 355 360 365

Please Note:

F4I:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/674,593

DATE: 07/05/2001
TIME: 15:41:03

Input Set : A:\538889_1.txt
Output Set: N:\CRF3\07032001\I674593.raw

L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:410 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11
L:411 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11
L:413 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11
L:414 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11
L:415 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11